

Conference Abstract

# The Additivity Project: Achieving additivity of structured taxonomic character data by persistently linking them to individual specimens

Patrick Plitzner<sup>‡</sup>, Tilo Henning<sup>‡</sup>, Andreas Müller<sup>‡</sup>, Anton Güntsch<sup>‡</sup>, Walter G. Berendsohn<sup>‡</sup>, Thomas Borsch<sup>‡</sup>, Norbert Kilian<sup>‡</sup>

<sup>‡</sup> Botanischer Garten Botanisches Museum, Freie Universität Berlin, Berlin, Germany

Corresponding author: Patrick Plitzner ([p.plitzner@bgbm.org](mailto:p.plitzner@bgbm.org))

Received: 12 Jun 2019 | Published: 18 Jun 2019

Citation: Plitzner P, Henning T, Müller A, Güntsch A, Berendsohn W, Borsch T, Kilian N (2019) The Additivity Project: Achieving additivity of structured taxonomic character data by persistently linking them to individual specimens. Biodiversity Information Science and Standards 3: e37178. <https://doi.org/10.3897/biss.3.37178>

## Abstract

Herbarium specimens have always played a central role in the classical disciplines of plant sciences and the global digitisation efforts now open new horizons. To make full use of the inherent possibilities of specimen based taxonomic descriptions corresponding workflows are needed. A crucial step in the comparative analyses of organisms is the preparation of a character matrix to record and compare the morphological variation of taxa on the basis of individual specimens.

This project focuses on the optimisation of the taxonomic research process with respect to delimitation and characterisation (“descriptions”) of taxa (Henning et al. 2018). The angiosperm order Caryophyllales provides exemplar use cases through cooperation with the Global Caryophyllales Initiative (Borsch et al. 2015). The workflow for sample data handling (Kilian et al. 2015), implemented on the EDIT Platform for Cybertaxonomy (<http://www.cybertaxonomy.org>, Ciardelli et al. 2009), has been extended to support additive characterisation of taxa via specimen character data.

The Common Data Model (CDM), already supporting persistent inter-linking of specimens and their metadata (Plitzner et al. 2017), has been adapted to facilitate specimen

descriptions with characters constructed from the combination of structure and property terms and their corresponding states. Semantic web technology is used to establish and continuously elaborate expert community-coordinated exemplar vocabularies with term ontologies and explanations for characters and states (GFBio Terminology Service, Karam et al. 2016). Character data are recorded and stored in structured form in character state matrices for individual specimens instead of taxa, which allows generation of taxon characterisations by aggregating the data sets for the individual specimens included. Separating characters in structures and properties, which are based on concepts in public ontologies, guarantees a high visibility and instant re-usability of these character data. Taking into account that taxon concepts evolve during the iterative knowledge generation process in systematic biology, additivity of character data from specimen to taxon level therefore greatly facilitates the construction and reproducibility of taxon characterisations from changing specimen and character data sets.

## Keywords

EDIT Platform for Cybertaxonomy, specimen, character data, semantic web, ontologies, taxon description

## Presenting author

Patrick Plitzner

## Presented at

Biodiversity\_Next 2019

## References

- Borsch T, Hernández-Ledesma P, Berendsohn W, Flores-Olvera H, Ochoterena H, Zuloaga F, von Mering S, Kilian N (2015) An integrative and dynamic approach for monographing species-rich plant groups – Building the global synthesis of the angiosperm order Caryophyllales. *Perspectives in Plant Ecology, Evolution and Systematics* 17 (4): 284-300. <https://doi.org/10.1016/j.ppees.2015.05.003>
- Ciardelli P, Kelbert P, Kohlbecker A, Hoffmann N, Güntsch A, Berendsohn WG (2009) The EDIT Platform for Cybertaxonomy and the taxonomic workflow: selected Components. *Lecture Notes in Informatics* 154: 625-638.
- Henning T, Plitzner P, Güntsch A, Berendsohn W, Müller A, Kilian N (2018) Building compatible and dynamic character matrices – Current and future use of specimen-based character data. *Botany Letters* 165: 352-360. <https://doi.org/10.1080/23818107.2018.1452791>

- Karam N, Müller-Birn C, Gleisberg M, Fichtmüller D, Tolksdorf R, Güntsch A (2016) A Terminology Service Supporting Semantic Annotation, Integration, Discovery and Analysis of Interdisciplinary Research Data. *Datenbank-Spektrum* 16 (3): 195-205. <https://doi.org/10.1007/s13222-016-0231-8>
- Kilian N, Henning T, Plitzner P, Müller A, Güntsch A, Stöver B, Müller K, Berendsohn W, Borsch T (2015) Sample data processing in an additive and reproducible taxonomic workflow by using character data persistently linked to preserved individual specimens. *Database* 2015 <https://doi.org/10.1093/database/bav094>
- Plitzner P, Müller A, Güntsch A, Berendsohn W, Kohlbecker A, Kilian N, Henning T, Stöver B (2017) The CDM Applied: Unit-Derivation, from Field Observations to DNA Sequences. *Proceedings of TDWG* 1 <https://doi.org/10.3897/tdwgproceedings.1.20366>